nature genetics

Distributions of epistasis in microbes fit predictions from a fitness landscape model

Guillaume Martin[1,2], Santiago F Elena[3] & Thomas Lenormand[1]

How do the fitness effects of several mutations combine? Despite its simplicity, this question is central to the understanding of multilocus evolution. Epistasis (the interaction between alleles at different loci), especially epistasis for fitness traits such as reproduction and survival, influences evolutionary predictions[1,2] "almost whenever multilocus genetics matters"[3]. Yet very few models[4,5] have sought to predict epistasis, and none has been empirically tested. Here we show that the distribution of epistasis can be predicted from the distribution of single mutation effects, based on a simple fitness landscape model[6]. We show that this prediction closely matches the empirical measures of epistasis that have been obtained for *Escherichia coli*[7] and the RNA virus vesicular stomatitis virus[8]. Our results suggest that a simple fitness landscape model may be sufficient to quantitatively capture the complex nature of gene interactions. This model may offer a simple and widely applicable alternative to complex metabolic network models, in particular for making evolutionary predictions.

Recent technical improvements in genetics have allowed the measurement of epistatic interactions in a very precise way. Large amounts of empirical evidence stemming from the study of development, metabolic networks[5] and quantitative traits analyses[9] have accumulated to show that epistasis is a widespread feature of genetic systems. However, despite many examples of epistatic interactions between particular pairs of loci, relatively little is known about the overall distribution of epistasis among random sets of mutations scattered across the genome. A few studies have sought to measure this distribution directly in model systems such as *Escherichia coli*[7], RNA viruses[8,10,11] and *Saccharomyces cerevisiae*[12]. These studies have shown that the variance of epistatic interactions is large compared with their mean, which is always relatively close to zero[10].

Unfortunately, theoretical developments have not gone hand-in-hand with those empirical advances, and in particular, no theory is yet available to explain, predict or generalize those observations. Fitness epistasis among mutations at enzymatic loci has been modeled with metabolic control theory[4] or flux balance analysis[5]. The former assumes idealized metabolic pathways and specific metabolism-fitness relationships, whereas the latter models a much more precise and complete metabolic network based on genomic data from model

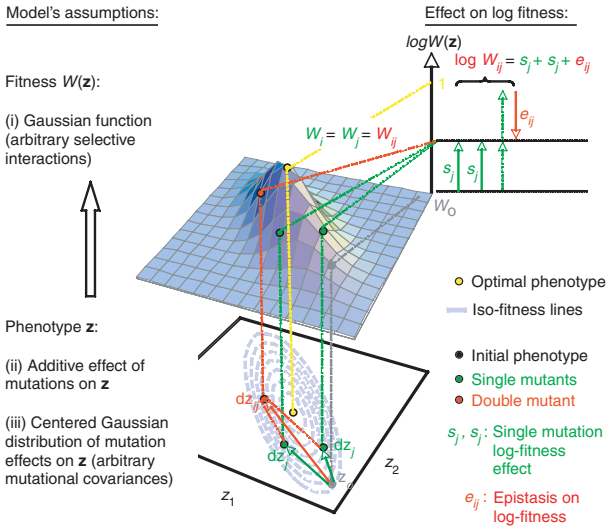

Figure 1 Fitness landscape model of epistasis between mutations, based on three main assumptions (i–iii). An example of the fitness landscape is given with only two phenotypic traits determining fitness and two mutations *i* and *j* (both beneficial here). Fitness $W(\mathbf{z})$ decreases as a multivariate Gaussian function of the distance to the optimum on both traits, with arbitrary interactions between traits (assumption (i)). From an arbitrary initial phenotype ($\mathbf{z}_o$), distinct mutations (at different loci) produce random perturbations of phenotypic traits ($\mathbf{dz}_i$,$\mathbf{dz}_j$), which act additively on phenotype when combined together ($\mathbf{dz}_{ij}$) (assumption (ii)). Although mutation effects on $\mathbf{z}$ are additive, epistasis measured on log relative fitness (e_{ij}) is naturally generated by the nonlinear mapping from phenotype $\mathbf{z}$ to fitness $W(\mathbf{z})$. With stabilizing selection, the curvature of $W(\mathbf{z})$ produces a diminishing return of fitness on phenotype so that two mutations whose effects 'add up' for phenotype do not 'add up' for fitness (here, the outcome is negative e_{ij}). The more precise quantitative predictions regarding e_{ij} distributions depend on the type of distribution chosen for the $\mathbf{dz}_i$ (in our model, a multivariate Gaussian with arbitrary mutational covariances (assumption (iii)).

[1]Centre d'écologie fonctionnelle et évolutive–Centre National de la Recherche Scientifique UMR 5175, 1919 Route de Mende, 34293 Montpellier, France. [2]Département d'Ecologie et Evolution, Bâtiment Biophore, Université de Lausanne, 1015 Switzerland. [3]Instituto de Biología Molecular y Celular de Plantas, Consejo Superior de Investigaciones Científicas-UPV, 46022 Valencia, Spain. Correspondence should be addressed to G.M. (guillaume.martin@unil.ch).

Received 6 September 2006; accepted 16 February 2007; published online 18 March 2007; doi:10.1038/ng1998

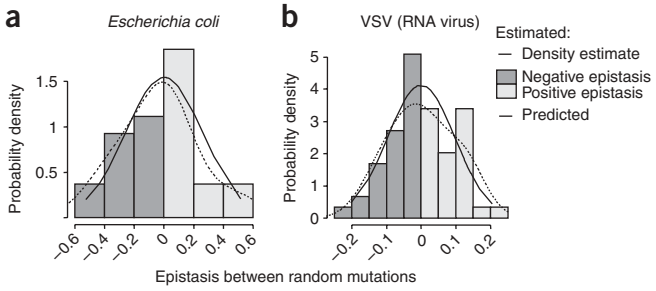

Figure 2 Observed and predicted distributions of fitness epistasis between random pairs of mutations. (**a**,**b**) The observed distribution of epistasis for log-fitness is presented for two model species (**a**, *E. coli* data set 1 (ref. 7) and **b**, VSV data set[8]), along with the predicted Gaussian distribution $N(0, 2v_s^*)$, where v_s^* is the variance of single fitness effects (at the optimum, $s_0 = 0$) estimated directly (*E. coli*) or inferred (VSV, using the correction for $s_0 \neq 0$; Methods). The dashed line gives the kernel density estimate of the data (a smoothed equivalent of a histogram) with a Gaussian smoothing kernel. The model and data are in very good agreement for both species.

organisms. These approaches propose a clear and valuable mechanistic basis for gene interaction, and flux balance analysis has even been successfully tested[13], although only for the extreme case of single-gene knockouts. However, flux balance analysis requires extensive knowledge of the metabolism of particular organisms in particular environments, so it cannot be applied to a wide range of biological systems. Perhaps more importantly, the predicted patterns of epistasis remain to be directly tested with empirical data.

Despite the relatively minor function that Fisher attributed to epistasis in adaptation, his geometrical model of adaptation[14] provides a general, yet unexplored, framework to predict epistasis among mutations. This model assumes stabilizing selection on '*n*' phenotypic traits. The effect of a mutation is modeled as a random displacement in this '*n*-dimensional' phenotypic space. Although it has been very useful in rejuvenating the theory of adaptation[15], Fisher's model is often merely viewed as a heuristic picture for mutational effects. However, by avoiding a mechanistic description of the relationship between particular mutations, phenotypes and fitness, it allows a global description of mutational effects without exhaustive knowledge of the underlying genetic details. This generality is what makes it attractive[16]. In addition, many of the underlying assumptions in Fisher's original model are in fact quite realistic[17] or can easily be relaxed[6,18].

Here we used an extended version of Fisher's geometric model[6] that allows for arbitrary mutational and selective interactions between traits determining fitness (**Fig. 1**). Any model of stabilizing selection (selection for a given optimum) naturally generates epistasis for fitness, even when mutations act additively on the underlying phenotype, because the relationship between phenotype and fitness is nonlinear, as in our model here (**Fig. 1**). The model is formulated in terms of measurable quantities (focusing on mutant fitness W instead of underlying phenotype $\mathbf{z}$ (**Fig. 1**), which makes it directly comparable to observation. We chose a Gaussian fitness function (relating phenotype to fitness) because it approximates any smooth function in the vicinity of an optimum and it qualitatively predicts observed patterns in empirical data, such as the gamma distribution of mutational effects in benign environments[6] and the effect of environmental harshness on the mutational mean and variance in fitness[19]. Finally, the model can be easily generalized to describe epistasis among more than two mutations (**Supplementary Methods** online).

From this model, we derived three testable predictions (for details and interpretation, see Methods). $\text{Log}(w_i)$ denotes the log-fitness of a mutant bearing mutation i relative to that of the nonmutated initial genotype (equation (1)). Epistasis among the pair of mutations i and j (e_{ij}) is defined (equation (2)) as the difference between the log-fitness of the double mutant and that expected if mutations acted multiplicatively: $e_{ij} = \log(w_{ij}) - \log(w_i\, w_j)$. The model first predicts that the probability density function of e_{ij} is well approximated by a Gaussian with mean zero and variance $2v_s^*$, where $v_s^* = Var(\log(w_i))$ is the variance of single mutation effects measured in an environment to which the initial genotype is well adapted (equation (3)). The second prediction of the model is that epistasis among pairs of beneficial mutations should be both biased and skewed toward negative values. Third, the model predicts that when the initial genotype is at or near the optimum, the distribution of log-fitness among mutant lines with k mutations (all deleterious) can be approximated by a gamma distribution $\Gamma(\beta, \alpha)$ with a constant shape (β) and a scale (α) proportional to k.

Table 1 Pairwise epistasis in VSV and *E. coli*: fit to predictions

Species (number of observations)	Input parameters (number of observations)	Epistasis variance $v_{e\ obs}$ (s.e.m.) $v_{e\ pred}$ (s.e.m.)	$v_{e\ obs}$ / $v_{e\ pred}$ *F* test	Epistasis mean $\mu_{e\ obs}$ (s.e.m.) *t* test	K-S test
E. coli ($n = 27$) (all mutations) (ref. 7)	$v_s^* = 0.033$ ($n = 54$)	$v_{e\ obs} = 0.0652$ (0.017) $v_{e\ pred} = 0.0670$ (0.013)	$F_{26,53} =$ **97%** $P = 0.94$	$\mu_{e\ obs} = -0.033$ (0.05) $\mu_{e\ pred} = 0$ $t_{26} = -0.66$, $P = 0.51$	$D = 0.18$ $P = 0.30$
VSV ($n = 59$) (all mutations) (ref. 8)	$v_s^* = 0.0047$ $s_0 = 0.11^a$ ($n = 118$)	$v_{e\ obs} = 0.0089$ (0.0016) $v_{e\ pred} = 0.0094$ (0.0012)	$F_{58,117} =$ **95%** $P = 0.86$	$\mu_{e\ obs} = 0.004$ (0.012) $\mu_{e\ pred} = 0$ $t_{58} = 0.31$, $P = 0.75$	$D = 0.085$ $P = 0.76$
VSV ($n = 15$) (beneficial mutations) (ref. 8)	$n_e = 3^a$ $\lambda_e = 0.06^a$ $s_0 = 0.11^a$ ($n = 118$)	$v_{e\ obs} = 0.0043$ (0.0016) $v_{e\ pred} = 0.0045$	$^{obs}/_{pred} =$ **96%** $\chi^2_{14} = 13.4$ $P = 0.99$	$\mu_{e\ obs} = -0.075$ (0.023) $\mu_{e\ pred} = -0.059$ $t_9 = -0.95$, $P = 0.36$	$D = 0.27$ $P = 0.22$

In this test of the fit to predictions, observed variance ($v_{e\ obs}$) and predicted variance ($v_{e\ pred}$) were compared with two-tailed *F* tests, as the prediction ($v_{e\ pred} = 2v_s^*$) is itself based on an independent estimate of v_s^*. For beneficial mutations in VSV (row three), the simulated prediction (see Methods) was considered exact (a conservative approach), so a two-tailed χ^2 test was used. Observed means ($\mu_{e\ obs}$) and predicted means ($\mu_{e\ pred}$) were compared with two-tailed *t* tests. Distributions were not significantly different from a Gaussian (Shapiro-Wilks test: $P = 0.45$, *E. coli*; $P = 0.44$, VSV (all mutations); $P = 0.59$, VSV (beneficial mutations)). The predicted and observed overall distributions were also compared with Kolmogorov-Smirnov (K-S) tests (one-sample test with $N(0, 2v_s^*)$ or two-sample test with the simulated prediction for beneficial mutations in VSV). Column two gives the value of the input parameters used in the prediction, estimated independently from the log-fitness distributions of the single mutants from which the double mutants were derived. None of the observed distributions differ significantly from the predictions. The ratios of variances (in boldface) show that the prediction is always within 5% of the estimation. Power curves for the *t* tests and *F* tests are given in **Supplementary Figure 1**.

[a]Estimates from the fit of the VSV data set (displaced gamma), with $n = n_e = 3$ in simulations (**Fig. 3**) closest integer to the estimated $n_e = 2.5$.

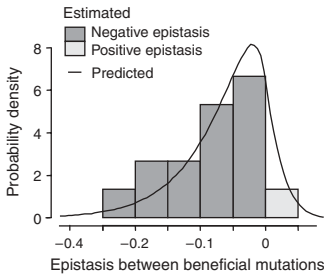

Figure 3 Observed and predicted distributions of epistasis between VSV beneficial mutations[8] (15 epistasis estimates). The predicted distribution (continuous line) is obtained by simulations calibrated with the estimates of n_e, s_o and λ_e from single mutant log-fitness values (see Methods and **Table 1**). We drew 1,500 mutant phenotypic effect vectors, $\mathbf{dx}_i$, into a standard multivariate Gaussian of dimension $n = n_e = 3$, which is the closest integer to $n_e = 2.5$ estimated from the VSV data. A vector of the multivariate distance to the optimum, $\mathbf{x}_o$, was drawn into the same distribution and scaled so that $-1/2\ \lambda_e \log(\mathbf{x}_o^T \cdot \mathbf{x}_o) = s_o$ (with $s_o = 0.11$ and $\lambda_e = 0.06$, from the data). The epistasis coefficient between pairs of mutations (i,j) was computed as $e_{ij} = -\lambda_e\, \mathbf{dx}_i^T \cdot \mathbf{dx}_j$, and we kept only the subset of simulated mutants with beneficial single effect (that is, $\mathbf{dx}_i$, such that $\log(w_i) = -\lambda_e(\mathbf{x}_o^T \cdot \mathbf{dx}_i + 1/2\mathbf{dx}_i^T \cdot \mathbf{dx}_i) > 0$); the resulting distribution is that of e_{ij} among all possible pairs of beneficial mutations for this simulation. This was repeated 20 times to account for the effect of variation of the direction of $\mathbf{x}_o$ for a given s_o. The predicted distribution, obtained by the overall distribution of e_{ij} among the 20 replicated simulations, is close to the empirical distribution (histogram).

These three predictions have been tested with data from two widely different species: *E. coli*[7] and vesicular stomatitis virus[8] (VSV; **Fig. 2**). For each species, we tested three predictions (equation (3)): among all pairs of mutations, $E(e_{ij}) = 0$, $Var(e_{ij}) = 2v_s^*$, and $e_{ij} \sim N(0,2v_s^*)$. None of these hypotheses was rejected. As already outlined in the original studies, $E(e_{ij})$ did not significantly depart from 0 in either *E. coli* or VSV. Furthermore, the observed $Var(e_{ij})$ values were very close to the predicted values (<5% difference) when synthetic lethals were discarded. Finally, overall, the distributions were not significantly different from the predicted Gaussian $N(0,2v_s^*)$ (**Table 1** and **Fig. 2**). The power curves for these tests (**Supplementary Fig. 1** online) also indicated that even if small departures from the null hypotheses could not be detected, the 'true' differences between observations and predictions were likely to be relatively small (except perhaps for the $E(e_{ij})$ value observed for *E. coli*).

We tested the second prediction with the subset of beneficial mutations analyzed in the VSV experiments. The model does not provide an analytical expression for the distribution of e_{ij} among beneficial mutations; therefore, we did this test by comparing the observed epistasis with a simulated distribution using parameters inferred from the distribution of single-mutant fitnesses (**Table 1**). We found good agreement between the empirical and predicted distributions (**Fig. 3**; among the subset of beneficial mutations, $E(e_{ij}) < 0$ is expected and observed). None of the predictions regarding the mean, variance or distribution of e_{ij} was rejected, and the power of these tests was reasonably good (**Supplementary Fig. 1**). However, this comparison has obvious limitations, as it was based on only 15 double mutants constructed from six distinct beneficial mutations.

Testing the third prediction required fitting gamma distributions to the log-fitness of mutant lines carrying a known number k of mutations to check whether the change in shape (β_k) and scale (α_k) with k conforms to the prediction. We used an extensive collection of *E. coli* genotypes differing in the location and number of transposon insertions[7] (see Methods). These data sets consisted of log-fitness measures of genotypes carrying either $k = 1$, 2 or 3 random insertions, with 75 different combinations per k value. To estimate the log-fitness distributions for each k value, we fitted independent gamma distributions to the data by maximum likelihood (model 1 of **Table 2**; see Methods), discarding four synthetic lethals. We compared these estimates with alternative data fits in which α and/or β were constrained according to k values, following alternative predictions. We determined the Akaike information criterion (AIC)[20] for several alternative models and estimated parameters for the best-fitting models (**Table 2**). The model following our theoretical expectation exactly (equation (A.5) of the **Supplementary Methods**: $\beta_k = \beta_1$ and $\alpha_k = k\alpha_1$) was the most adequate (lowest AIC, model 4). When directly estimated (model 1), the parameters were indeed very close to the model prediction (**Fig. 4** and **Supplementary Table 1** online). Thus, the form of epistasis and its impact on log-fitness distributions seems accurately captured by the model.

Overall, the distribution of epistasis for fitness predicted from a simple fitness landscape model adequately accounted for empirical distributions among both pairs and triplets of nonlethal mutations (including between beneficial ones). To our knowledge, this is the first empirical support for a general model of epistasis. The inherent simplicity of this model allows testable predictions that are rarely available at this degree of generality in other models connecting genotype to phenotype to fitness, and it offers an alternative to the more complex and specific metabolic network models that have been developed[5]. The quantitative accuracy and generality of the predictions, which are independent of the adaptation level, number of traits or phenotypic correlations, could be useful in many evolutionary predictions (such as the evolution of sex[21,22]). Finally, the fit between this general model and data from two very different species indicates that the observed empirical patterns may apply to other species as well. In particular, among a set of random mutations, the average epistasis is small (close to zero) with a large variance (twice that of single effects at the optimum), whereas among only beneficial mutations, average epistasis is negative. The generality of these patterns remains an empirical issue, and, being fairly quantitative, they are open to further testing, particularly in higher eukaryotes.

Table 2 Effect of the number of mutations in *E. coli*: model comparison

Model	Constraints	d.f.	Dev	AIC
1[a]	$\alpha_1, \alpha_2, \alpha_3 \mid \beta_1, \beta_2, \beta_3$	8	0	16
2	$\alpha_1, \alpha_2, \alpha_3 \mid \beta_1 = \beta_2 = \beta_3$	6	0.08	12.08
3	$\alpha_1 = \alpha_2 = \alpha_3 \mid \beta_1, \beta_2, \beta_3$	6	8.24	20.24
4[b]	$\alpha_2 = 2\alpha_1, \alpha_3 = 3\alpha_1 \mid \beta_1 = \beta_2 = \beta_3$	4	2.32	**10.32**
5	$\alpha_1 = \alpha_2 = \alpha_3 \mid \beta_1 = \beta_2 = \beta_3$	4	21.8	29.8

Models fitted for *E. coli* data set 2 (ref. 7). β_k and α_k refer to the shape and scale of the gamma distributions fitted to the log-fitness distributions of distinct mutant sets (carrying $k = 1$, 2 or 3 mini-Tn*10* insertions). Models are spelled out by the constraints imposed among these parameters in the fitting process. Each model includes the number of fitted parameters (d.f.), the residual deviance (Dev) and the AIC. Model 1 is unconstrained and provides the direct estimation of β_k and α_k; the distributions for $k = 1$, 2 or 3 are fitted independently. Its fit is not significantly better ($\chi^2 = 2.32$, 4 d.f., $P = 0.68$) than that of model 4 (more constrained), which corresponds to our prediction (constant shape β and increasing scale $\alpha_k = k\alpha_1$; **Supplementary Methods**) and is the most adequate model (lowest AIC, in boldface). Similarly, model 2 (constraining shapes to be identical) is not significantly better than model 4 ($\chi^2 = 2.24$, 2 d.f., $P = 0.33$). Model 3, corresponding to multiplicative fitness effects ($e_{ij} = 0$; **Supplementary Methods**), and model 5 (all distributions equal) are both rejected, showing that the analysis has sufficient statistical power to reject inaccurate models. [a]Estimate: shape β and scale α fitted independently for each mutant subsets. [b]Our prediction: constant shape β and increasing scale $\alpha_k = k\alpha_1$.

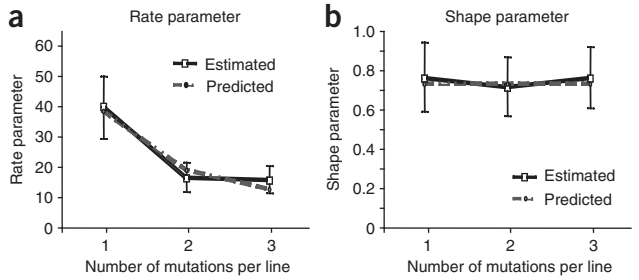

Figure 4 Observed and predicted change of the distribution of log-fitness with the number of mini-Tn*10* insertions in *E. coli*: gamma approximation. Parameters of a gamma distribution (its rate $1/\alpha$ (**a**) and shape β (**b**)) estimated by maximum likelihood on *E. coli* data set 2 (ref. 7) for each subset of mutants (single, double and triple). The estimated values were obtained by independent fitting of gamma distributions on each subset (model 1, **Table 2**). The predicted values were obtained by constraining the parameters of the gamma across subsets to follow our prediction (model 4, **Table 2**), with α proportional to the number of mutations per line in each subset (one, two or three) and constant β across subsets. Bars give the support limits for the estimates (**Supplementary Table 1**). The unconstrained fit is not significantly better than the fit obtained under our prediction (model 4 versus model 1, $\chi^2 = 2.32$, 4 d.f., $P = 0.68$).

METHODS

Model. Following Lande[23], the fitness $W(\mathbf{z})$ of a phenotype $\mathbf{z}$ (of dimension n, the number of phenotypic traits under selection) is given by a multivariate Gaussian function $W(\mathbf{z}) \equiv \exp(-1/2\ \mathbf{z}^{T}\cdot\mathbf{S}\cdot\mathbf{z})$ where ^{T} represents transposition, and $\mathbf{S}$ is an arbitrary $n \times n$ symmetric positive semi-definite matrix that describes all the selective interactions between phenotypic traits. Therefore, the fitness function may be flat on some directions of the landscape (if $\mathbf{S}$ has some null eigenvalues). The phenotype vector of the initial nonmutated genotype is $\mathbf{z}_o$, and its level of adaptation is measured by $s_o \equiv -\log W(\mathbf{z}_o) = 1/2\ \mathbf{z}_o^{T}\cdot\mathbf{S}\cdot\mathbf{z}_o$. The phenotypic effect of a mutation i is given by the vector $\mathbf{dz}_i$, so the mutant phenotype is $\mathbf{z}_o + \mathbf{dz}_i$, and its log-relative fitness is as follows[6]:

$$\log(w_i) = \log(W(\mathbf{dz}_i+\mathbf{z}_o)/W(\mathbf{z}_o)) = -\mathbf{z}_o^{T}\cdot\mathbf{S}\cdot\mathbf{dz}_i - {}^{1}\!/_{2}\mathbf{dz}_i^{T}\cdot\mathbf{S}\cdot\mathbf{dz}_i \quad (1)$$

Here, $\mathbf{dz}_i$ refers to genotypic values (averaged over replicates of a given line), not to individual replicate phenotypes (which are also influenced by micro-environmental effects). It can be shown[24] (**Supplementary Note** online) that if fitness $W(\mathbf{z})$ is a Gaussian function of individual phenotypes, then log-relative fitness is necessarily a quadratic function of genotypic values, as used here, so environmental effects on phenotype can be ignored. This model has received indirect support from empirical distributions of single-mutation effects[6,19]. The effects of mutations on phenotype were assumed to be additive, so that the joint effect of mutations i and j is $\mathbf{dz}_{ij} = \mathbf{dz}_i + \mathbf{dz}_j$. Notably, the model is robust to this fairly strong assumption (**Supplementary Fig. 2** online). From equation (1), if $w_{ij} = W(\mathbf{z}_o + \mathbf{dz}_{ij}) / W(\mathbf{z}_o)$ is the relative fitness of the double mutant, then pairwise epistasis, defined as $e_{ij} = \log(w_{ij}) - \log(w_i\ w_j)$, is given by the following:

$$e_{ij} = -\mathbf{dz}_i^{T}\cdot\mathbf{S}\cdot\mathbf{dz}_j \quad (2)$$

From equation (2), it seems that fitness epistasis between two given mutations does not depend on $\mathbf{z}_o$, the position of the wild-type in phenotypic space (that is, of its degree of adaptation to the environment). This stems from the fact that we assumed a Gaussian fitness function. Indeed, epistasis e_{ij} depends on the curvature of the log-fitness function ($\log W(\mathbf{z})$; **Fig. 1**), and this curvature is the same for all $\mathbf{z}$ values with a quadratic log-fitness (or Gaussian fitness).

To predict the distribution of e_{ij} from equation (2), assumptions must be made about the distribution of mutation effects on phenotype ($\mathbf{dz}$). We assume that $\mathbf{dz}$ is drawn into a multivariate Gaussian distribution with mean $\mathbf{0}$ and arbitrary covariance matrix $\mathbf{M}$. Then, e_{ij} is a bilinear form in Gaussian vectors[25] whose moments can be related to those of the distribution of single effects, $\log(w_i)$ in equation (1). At the optimum ($s_o = 0$ and $\mathbf{z}_o = \mathbf{0}$ in equation (1)), the variance v_s^* of $\log(w_i)$ is[6] $v_s^* = Var(-1/2\ \mathbf{dz}_i^{T}\cdot\mathbf{S}\cdot\mathbf{dz}_i) = 1/2\ \mathrm{Tr}((\mathbf{S}\cdot\mathbf{M})^2)$, where 'Tr(·)' is matrix trace. Now, from equation (2) and because $\mathbf{dz}_i$ and $\mathbf{dz}_j$ are independent ($Cov(\mathbf{dz}_i,\mathbf{dz}_j) = \mathbf{0}$), we find for e_{ij} a mean $\mu_e \equiv E(e_{ij}) = E(-\mathbf{dz}_i^{T}\cdot\mathbf{S}\cdot\mathbf{dz}_j) = 0$, and because the phenotypic effects $\mathbf{dz}_i$ are multivariate Gaussian, the variance of epistasis is given by the following:

$$v_e \equiv Var(e_{ij}) = Var(-\mathbf{dz}_i^{T}\cdot\mathbf{S}\cdot\mathbf{dz}_j) = \mathrm{Tr}((\mathbf{S}\cdot\mathbf{M})^2) = 2v_s^* \quad (3)$$

For the same reasons, the distribution of e_{ij} has also no 'skewness', so the Gaussian $N(0,\ 2v_s^*)$ provides a simple and accurate approximation to the distribution of e_{ij} (**Supplementary Fig. 3** online), depending on a single parameter (v_s^*). When the initial genotype is not at the optimum ($s_o > 0$), the variance of $\log(w_i)$ is not equal to v_s^* but must be corrected[6] with a measure of s_o, as $v_s^* = Var(\log(w_i))\ /\ (1 + 2s_o\ /\ \bar{s})$, where $\bar{s} = E(\log(w_i))$ is the mean effect of single mutations on log relative fitness. This correction was used for the VSV data set.

The prediction $v_e = 2\ v_s^*$ in equation (3) is fairly general. In particular, as equation (2) does not depend on $\mathbf{z}_o$, this relation is valid for any distance to the optimum s_o. Similarly, as equation (3) is valid for any $\mathbf{M}$ and $\mathbf{S}$, it does not depend on the details of the phenotypic landscape: the number of traits n (dimension of $\mathbf{M}$ and $\mathbf{S}$) or their mutational and selective correlations (elements of $\mathbf{M}$ and $\mathbf{S}$). Although these parameters do influence distributions of single-mutation fitness effects[6], they do not alter the relationship between single and multiple mutant fitness values.

Similarly, the distribution of e_{ij} among only beneficial mutations can be obtained for a given fitness landscape (although only by numerical simulations). This landscape is well characterized[6] by the 'effective trait effect' λ_e (which depends on the distribution of the eigenvalues of $\mathbf{S}\cdot\mathbf{M}$) and the 'effective number of dimensions' n_e. The distribution of e_{ij} (equation (2)) in the original landscape ($\mathbf{S}$, $\mathbf{M}$) is well approximated by its corresponding distribution in the equivalent landscape, for which $\mathbf{S} = \lambda_e\mathbf{I}_{ne}$ and $\mathbf{M} = \mathbf{I}_{ne}$, where $\mathbf{I}_{ne}$ is the identity matrix of dimension n_e. As for single-mutation effects[6], the approximation fits the two first moments of the whole distribution of e_{ij} among all mutations, but it is also accurate for the subset of beneficial ones (as shown by simulations in **Supplementary Fig. 3**). As both λ_e and n_e can be estimated from a data set of single mutation effects[6], e_{ij} among beneficial mutations can be numerically simulated for the same data set, as we did for the VSV data set (**Fig. 3**).

Epistasis among beneficial mutations tends to be negatively biased in this fitness landscape. This fact stems from the nonlinear relationship between phenotype and fitness. In the phenotypic landscape described in **Figure 1**, two beneficial mutations necessarily point to very similar directions (toward the optimum). However, because the fitness function is concave, there is a diminishing return of W on $\mathbf{z}$, along this direction. Therefore, two steps toward the optimum result in a lower fitness than would be expected from the addition of the fitness effects of each mutation, which corresponds to negative epistasis. More precisely, we consider the expression of e_{ij} as given in equation (2) but in the equivalent landscape (where $\mathbf{S} = \lambda_e\mathbf{I}_{ne}$). If $\mathbf{dx}_i$ and $\mathbf{dx}_j$ are the vectors of effects of mutations i and j in the equivalent landscape, then from equation (2), e_{ij} can be written as follows: $e_{ij} = -\lambda_e\ \mathbf{dx}_i^{T}\mathbf{dx}_j$. Therefore, the sign of e_{ij} is the opposite sign of the cosine of the angle between $\mathbf{dx}_i$ and $\mathbf{dx}_j$. Because two beneficial mutations tend to point to a similar direction (the optimum), their angle is small, so their cosine is positive. For this reason, e_{ij} between beneficial mutations tends to be negatively biased. Furthermore, this bias increases when the distance to the optimum decreases (simulations not shown). This is simply because the possible directions of $\mathbf{dz}$ resulting in beneficial mutations are more constrained when closer to the optimum, which results in a smaller angle between these mutations and therefore more negative epistasis.

The model can also easily be extended to lines carrying more than two mutations (**Supplementary Methods**). When $s_o = 0$, the distribution of log-fitness values among lines carrying k mutations has mean $E(\log(w|k)) = -k\bar{s}$ and variance $Var(\log(w|k)) = k^2v_s^*$, where $\bar{s}$ and v_s^* are the mean and variance of log-fitness values among single mutants ($k = 1$), when $s_o = 0$. Using the same approach reported before[6], the log-fitness distributions were approximated by a negative gamma matching the two first moments of the exact distribution given above. If we call α_k and β_k the scale and shape parameters of the gamma approximation for $\log(w|k)$, we then obtain $\alpha_k = k\alpha_1$ and $\beta_k = \beta_1$.

This simple approximation has good accuracy, as shown by the simulations in **Supplementary Figure 3**. In the absence of epistasis (additive log-fitness effects), the alternative prediction is given by $\alpha_k = \alpha_1$ and $\beta_k = k\beta_1$ (**Supplementary Methods**).

All our predictions stem from three necessary and sufficient assumptions (**Fig. 1**): (i) a Gaussian fitness function $W(\mathbf{z})$, (ii) additivity of the effects of mutations on phenotype ($\mathbf{z}$) and (iii) Gaussian distribution of these mutation effects with no bias (zero mean). The first two assumptions lead to equation (2), whereas the remaining predictions also require the third assumption to be valid but are fairly robust to the 'additivity' assumption, as shown in **Supplementary Figure 2**. The third assumption is overly restrictive: as the definition of traits is arbitrary, a transformation of 'real' phenotypic traits must exist such that mutational effects on these 'transformed' traits can be approximated by a Gaussian[6].

Data. To test our predictions, we chose the largest available epistasis data sets. However, we discarded studies based on standing genetic variation instead of newly arisen mutations, including a very large study measuring the whole distribution of epistasis in human immunodeficiency virus 1 (ref. 11). Indeed, the model is meant to describe newly arisen mutations, and its assumptions (such as symmetrical phenotypic distributions) may be invalidated by past selection altering the distribution of phenotypes. Several studies have measured (more or less directly) the mean of e_{ij} among random mutations in model species such as *Drosophila melanogaster*[26], yeast[12] or RNA viruses[10]; many found either a small or no departure from our expectation that $E(e_{ij}) = 0$ (refs. 3,10). However, fewer studies have directly measured its distribution by estimating individual epistasis coefficients among individual pairs of random mutations (not only those conferring visible phenotypes) of known effect and in the homozygous or haploid state (to avoid the confounding effect of dominance), which is required to fully test the predictions. To our knowledge, the data sets presented below are the only ones doing so and including enough replicates to ensure adequate power of the statistical analyses (**Supplementary Fig. 1**).

E. coli data set 1. This data set[7] (Fig. 3 of ref. 7) consists of fitness estimates in *E. coli* genotypes bearing one or two transposons ('mini-Tn*10*') insertions. Log relative fitness was estimated for the 27 possible pairs constructed from nine distinct bacterial genotypes bearing a single mini-Tn*10* insert with known fitness effect, yielding a complete set of $\log(w_i)$ and $\log(w_{ij})$ estimates. In this data set, the variance of single fitness effects v_s^* is directly available, as the initial genotype was well adapted to the test environment. This initial genotype has indeed evolved for 10,000 generations in this environment[27], and, correspondingly, not a single beneficial mutation has been detected among 225 insertions tested in that same environment[28]. Therefore, v_s^* was directly measured from $Var(\log(w_i))$ among the whole set of single mutants used to produce the double mutants.

E. coli data set 2. This data set[7] (Fig. 2 of ref. 7) corresponds to fitness estimates (replicated three times) for genotypes bearing $k = 1$, 2 or 3 random mini-Tn*10* inserts, with 75 distinct mutant genotypes per k value, and a corresponding fitness estimate of the non-mutated genotype (replicated 195 times). It is the largest data set available for the effect of random mutations, in combination or isolated, that we found in the literature. For the same reason outlined for the VSV data set, we discarded the nonviable genotypes from our analysis (one for $k = 2$ and three for $k = 3$).

VSV data set. In this data set, single and double mutants were created on an infectious cDNA by site-directed mutagenesis and were recovered after transfection of susceptible cells with the mutant plasmids[8]. As discussed[29], this recombinant virus was not well adapted to the environment in which measurements were taken, resulting in a proportion ($\approx 5\%$) of mutations with beneficial effects (of up to $\log(w_i) = 0.095$), so that $s_o > 0$. s_o was measured by maximum-likelihood fitting of a displaced gamma distribution[6] on the distribution of $\log(w_i)$ of the mutations used to construct the double mutants. The resulting estimate of $s_o = 0.11$ (s.e.m. = 0.01) was used to infer $v_s^* = Var(\log(w_i)) / (1 + 2s_o / \bar{s})$, where $\bar{s} = E(\log(w_i))$ is the mean fitness effect[6] (discussed above). The estimated shape $\beta = 1.94$ (s.e.m. = 0.47) and the scale $\alpha = 0.095$ (s.e.m. = 0.02) of this distribution were used to infer $n_e = 2.5$ and $\lambda_e = 0.06$ (using equations 6.a and 6.b of ref. 6), which were needed (along with s_o) to numerically predict the distribution of e_{ij} among beneficial mutations for this data set. Among the 62 double mutants studied, 3 were nonviable (synthetic lethal mutations; $w_{ij} = 0$) and were removed from the data set, as the model cannot account for lethal mutations. Also, the previously published results[8] are presented separately for pairs of beneficial and deleterious mutations (15 and 44 pairs, respectively, synthetic lethal mutations excluded), whereas the whole set of random mutations (a total of 44 + 15 = 59 pairs) must be used to test the first prediction presented above (**Fig. 2** and **Table 1**).

Different methods were used in each species to produce mutations (transposons for *E. coli* versus point mutations for VSV), but reviews of the empirical literature suggest that this should not result in strong differences for mutation effects on phenotype[30] or fitness traits[6]. The model we have presented here depends directly on phenotypic effects of mutations, not on their genetic nature, which may explain why it seems to accurately account for both the VSV and *E. coli* data sets.

Statistical analyses. To test the third prediction on VSV data set 2, three negative gamma distributions were simultaneously adjusted to the relative fitness values of single, double and triple mutants, with the variance in fitness estimated per line (from three independent replicate measures) taken into account. We assumed measurement error to be normally distributed, so that the fitness of each mutant j was drawn from $N(\mu_j, \sigma_j)$. We also assumed that μ_j values were drawn from a negative gamma distribution $\Gamma(\beta_k, \alpha_k)$, where $k = 1$, 2 or 3 refers to single, double or triple mutants, respectively. For raw data, there is a strong dependence between measurement error and the magnitude of fitness effects (the three repeated fitness measures of strongly deleterious mutants tend to show much larger variance). To account for this strong heteroscedasticity, we assumed a linear dependence of σ on μ_j. More specifically, σ_i was modeled as $\sigma + a\mu_j$, where a and σ are estimated from the data. Then, the likelihood of the data given the parameters (α_1, α_2, α_3, β_1, β_2, β_3, σ, a) was maximized according to the likelihood function in equation (A.6) of **Supplementary Methods**. Alternative models were constructed by constraining α_j and β_j values (**Table 2**, models 1–5). Models were compared based on their AIC score[20] and/or by likelihood ratio tests. This analysis confirmed the existence of a strong dependency of measurement error on the magnitude of fitness effects (for example, in model 1, $a = 0.21$, support limits {0.19, 0.25}).

Note: Supplementary information is available on the Nature Genetics website.

ACKNOWLEDGMENTS

We thank D. Waxman and P. Jarne for comments on this work. G.M. thanks J. Goudet for hosting him during part of this work. This work was supported by an Action Concertée Incitative from the French Ministry of Research (T.L.), a PhD fellowship from the French Ministry of Research (G.M.), the Swiss National Science Foundation (grant 31-108194/1 to G.M.) and the Spanish Ministerio de Educación y Ciencia (MEC)-FEDER grant BMC2003-00066 to S.F.E.

AUTHOR CONTRIBUTIONS

G.M. and T.L. designed the model, did the analysis and wrote the paper. S.F.E. compiled the data and wrote the paper.

COMPETING INTERESTS STATEMENT

The authors declare no competing financial interests.

Published online at http://www.nature.com/naturegenetics
Reprints and permissions information is available online at http://npg.nature.com/reprintsandpermissions

1. Phillips, P., Otto, S.P. & Whitlock, M.C. in *Epistasis and the Evolutionary Process* (eds. Wolf, J.B., Brodie, E.D. & Wade, M.J.) 20–38 (Oxford University Press, Oxford, 2000).
2. Otto, S.P. & Lenormand, T. Resolving the paradox of sex and recombination. *Nat. Rev. Genet.* **3**, 252–261 (2002).
3. Michalakis, Y. & Roze, D. Epistasis in RNA viruses. *Science* **306**, 1492–1493 (2004).
4. Szàthmary, E. Do deleterious mutations act synergistically - metabolic control-theory provides a partial answer. *Genetics* **133**, 127–132 (1993).
5. Segré, D., DeLuna, A., Church, G.M. & Kishony, R. Modular epistasis in yeast metabolism. *Nat. Genet.* **37**, 77–83 (2005).
6. Martin, G. & Lenormand, T. A multivariate extension of Fisher's geometrical model and the distribution of mutation fitness effects across species. *Evolution* **60**, 893–907 (2006).

7. Elena, S.F. & Lenski, R.E. Test of synergistic interactions among deleterious mutations in bacteria. *Nature* **390**, 395–398 (1997).
8. Sanjuàn, R., Moya, A. & Elena, S.F. The contribution of epistasis to the architecture of fitness in an RNA virus. *Proc. Natl. Acad. Sci. USA* **101**, 15376–15379 (2004).
9. Malmberg, R.L. & Mauricio, R. QTL-based evidence for the role of epistasis in evolution. *Genet. Res.* **86**, 89–95 (2005).
10. Burch, C.L., Turner, P.E. & Hanley, K.A. Patterns of epistasis in RNA viruses: a review of the evidence from vaccine design. *J. Evol. Biol.* **16**, 1223–1235 (2003).
11. Bonhoeffer, S., Chappey, C., Parkin, N.T., Whitcomb, J.M. & Petropoulos, C.J. Evidence for positive epistasis in HIV-1. *Science* **306**, 1547–1550 (2004).
12. Wloch, D.M., Borts, R.H. & Korona, R. Epistatic interactions of spontaneous mutations in haploid strains of the yeast *Saccharomyces cerevisiae*. *J. Evol. Biol.* **14**, 310–316 (2001).
13. Fong, S.S. & Palsson, B.O. Metabolic gene-deletion strains of *Escherichia coli* evolve to computationally predicted growth phenotypes. *Nat. Genet.* **36**, 1056–1058 (2004).
14. Fisher, R.A. *The Genetical Theory of Natural Selection,* (Oxford University Press, Oxford, 1930).
15. Orr, H.A. The genetic theory of adaptation: a brief history. *Nat. Rev. Genet.* **6**, 119–127 (2005).
16. Barton, N.H. & Keightley, P.D. Understanding quantitative genetic variation. *Nat. Rev. Genet.* **3**, 11–21 (2002).
17. Orr, H.A. The "sizes" of mutations fixed in phenotypic evolution: a response to Clarke and Arthur. *Evol. Dev.* **3**, 121–123 (2001).
18. Waxman, D. & Welch, J.J. Fisher's microscope and Haldane's ellipse. *Am. Nat.* **166**, 447–457 (2005).
19. Martin, G. & Lenormand, T. The fitness effect of mutations in stressful environments: a survey in the light of fitness landscape models. *Evolution* **60**, 2413–2427 (2006).
20. Akaike, H. A new look at the statistical model identification. *IEEE Trans. Automatic Control* **19**, 716–723 (1974).
21. Otto, S.P. & Feldman, M.W. Deleterious mutations, variable epistatic interactions and the evolution of recombination. *Theor. Popul. Biol.* **51**, 134–147 (1997).
22. Kouyos, R.D., Otto, S.P. & Bonhoeffer, S. Effect of varying epistasis on the evolution of recombination. *Genetics* **173**, 589–597 (2006).
23. Lande, R. The genetic covariance between characters maintained by pleiotropic mutations. *Genetics* **94**, 203–215 (1980).
24. Burger, R. in *The Mathematical Theory of Selection, Recombination, Mutation* Ch. V, 158–160 (John Wiley & Sons, Chichester, UK, 2000).
25. Mathai, A.M. & Provost, S.B. *Quadratic Forms in Random Variables* (Marcel Dekker, New York, 1992).
26. Whitlock, M.C. & Bourguet, D. Factors affecting the genetic load in *Drosophila*: synergistic epistasis and correlations among fitness components. *Evolution* **54**, 1654–1660 (2000).
27. Lenski, R.E. & Travisano, M. Dynamics of adaptation and diversification - a 10,000-generation experiment with bacterial-populations. *Proc. Natl. Acad. Sci. USA* **91**, 6808–6814 (1994).
28. Elena, S.F., Ekunwe, L., Hajela, N., Oden, S.A. & Lenski, R.E. Distribution of fitness effects caused by random insertion mutations in *Escherichia coli*. *Genetica* **103**, 349–358 (1998).
29. Sanjuàn, R., Moya, A. & Elena, S.F. The distribution of fitness effects caused by single-nucleotide substitutions in an RNA virus. *Proc. Natl. Acad. Sci. USA* **101**, 8396–8401 (2004).
30. Kidwell, M.G. & Lisch, D. Transposable elements as sources of variation in animals and plants. *Proc. Natl. Acad. Sci. USA* **94**, 7704–7711 (1997).